Knowledge Distillation
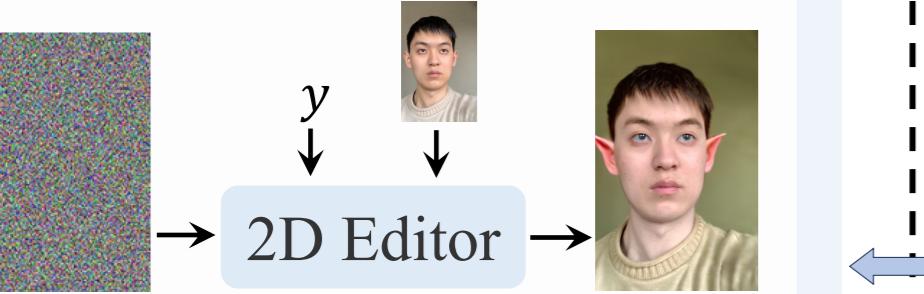
1. DDIM Inference

y

2D Editor
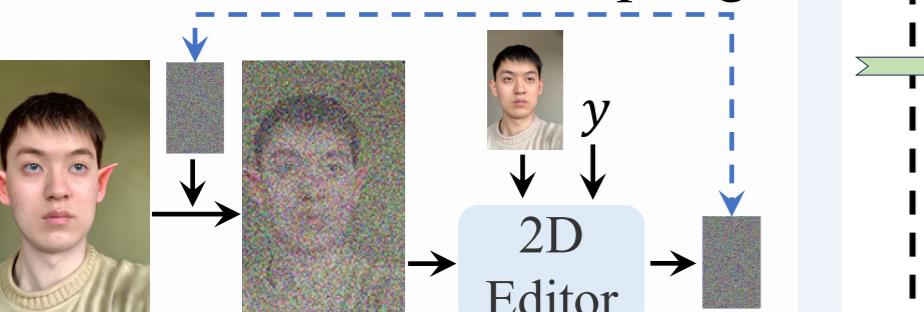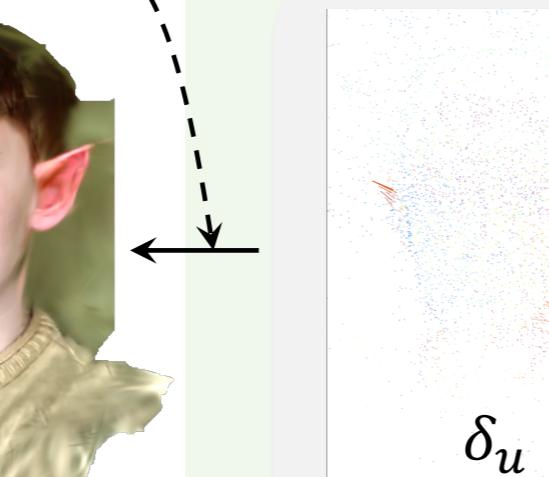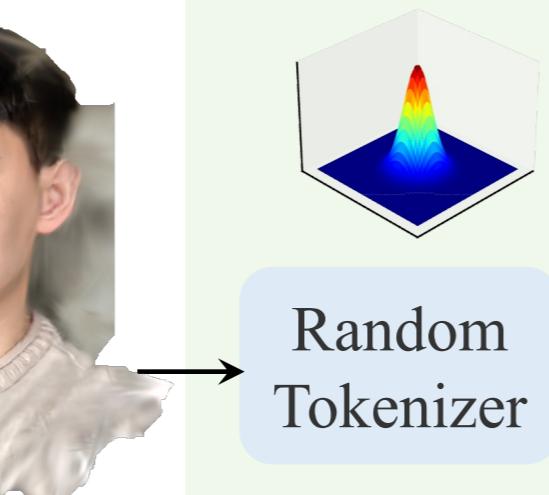
2. Diffusion Inversion

y

Denoising

3. Score Distillation Sampling

y

2D Editor

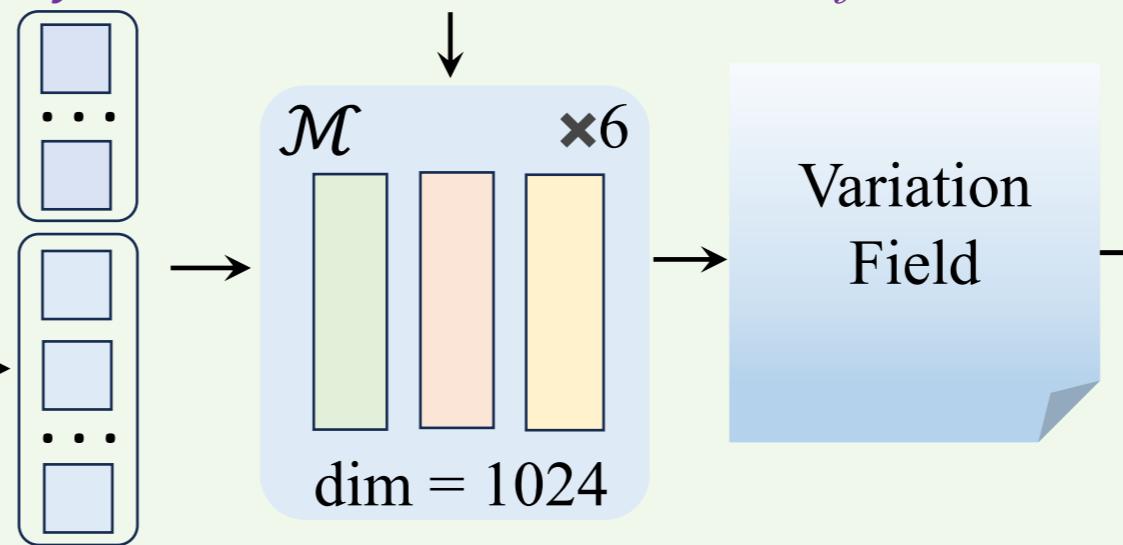Input ($\mathcal{X}^s$)
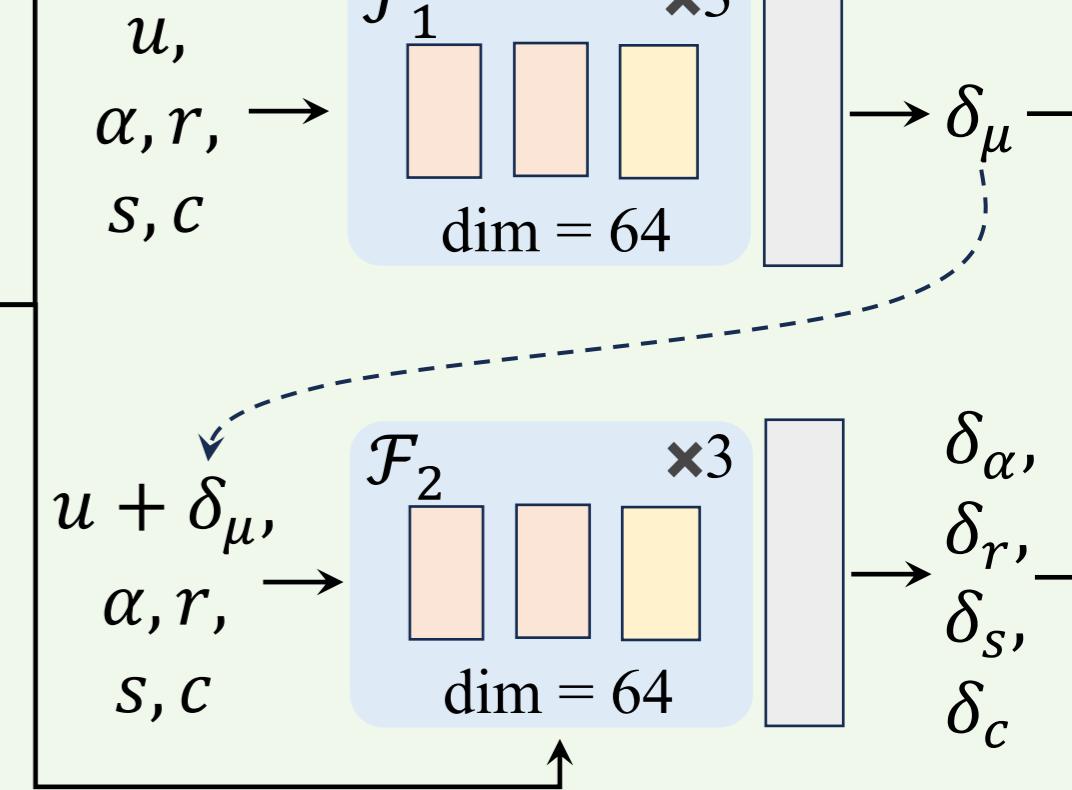
Output ($\mathcal{X}^r$)

Variation Predictor ($\mathcal{P}_\theta$)

y: *"Turn him into the Tolkien Elf"*

Random Tokenizer

$\mathcal{M}$ ×6

dim = 1024

Variation Field

$u,$ $\alpha, r,$ s, c

$\mathcal{F}_1$ ×3

dim = 64

δ_μ

$u + \delta_\mu,$ $\alpha, r,$ s, c

$\mathcal{F}_2$ ×3

dim = 64

$\delta_\alpha,$ $\delta_r,$ $\delta_s,$ δ_c

Visualization of the Variation

δ_u δ_α δ_s δ_c δ_r

Self-Attention   Cross-Attention

Feed Forward   Zero Linear

Loss Calculation   Gradient